

AD618748

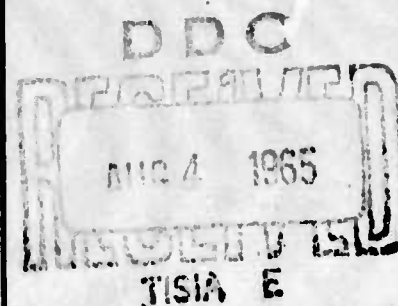
ORC 65-10  
APRIL 1965

# OPTIMIZATION IN OPERATIONS RESEARCH

by  
G. B. Dantzig

COPY	OF	<i>as</i>
HARD COPY		\$ . 1.00
MICROFICHE		\$ . 0.50

*11-p*



OPERATIONS RESEARCH CENTER

COLLEGE OF ENGINEERING

ARCHIVE COPY

UNIVERSITY OF CALIFORNIA - BERKELEY

OPTIMIZATION IN OPERATIONS RESEARCH: SOME EXAMPLES

by

George B. Dantzig  
Operations Research Center  
University of California, Berkeley

April 1965

ORC 65-10

This research has been partially supported by the Office of Naval Research under Contracts Nonr-222(83) and Nonr-3656(02) and the National Science Foundation under Grant GP 2633 and the U.S. Army Research Office under Contract DA-31-124-ARO-D-331 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

## OPTIMIZATION IN OPERATIONS RESEARCH: SOME EXAMPLES

by

George B. Dantzig

Mathematical methods are penetrating into everyday decision processes. Historically, planning has always been carried out using what some call "mature judgment". The human mind has a remarkable faculty for trying to consider many possible alternatives. The electronic computer with its equally remarkable speed, appears to be a natural tool for exploring alternative courses of action. It takes little imagination to see how computers can be programmed in the finest minutia to represent, say, the quantities of inventory on hand and the movement of in-process inventory from machine to machine in a job-shop. This obvious approach, called simulation, has popular appeal. Every manager, aware that he must keep up with the times, can easily imagine a computer sifting through the possible alternative combinations and selecting one that is best according to some criterion. In fact, the less such a person knows about mathematics, the more this simulation approach has appeal.

However, a raw search often can be combinatorially prohibitive. Consider, for example, a problem represented by a linear program. Mathematically, the set of possible feasible programs is represented by the intersection of a number of half-spaces whose boundaries are hyperplanes. The optimal solution consists in finding that vertex which is the greatest distance from a given objective-hyper-plane (the dotted one in Figure 1). The set of "feasible" solutions, (the shaded area) can be represented, of course, by linear combinations of vertices or extreme point

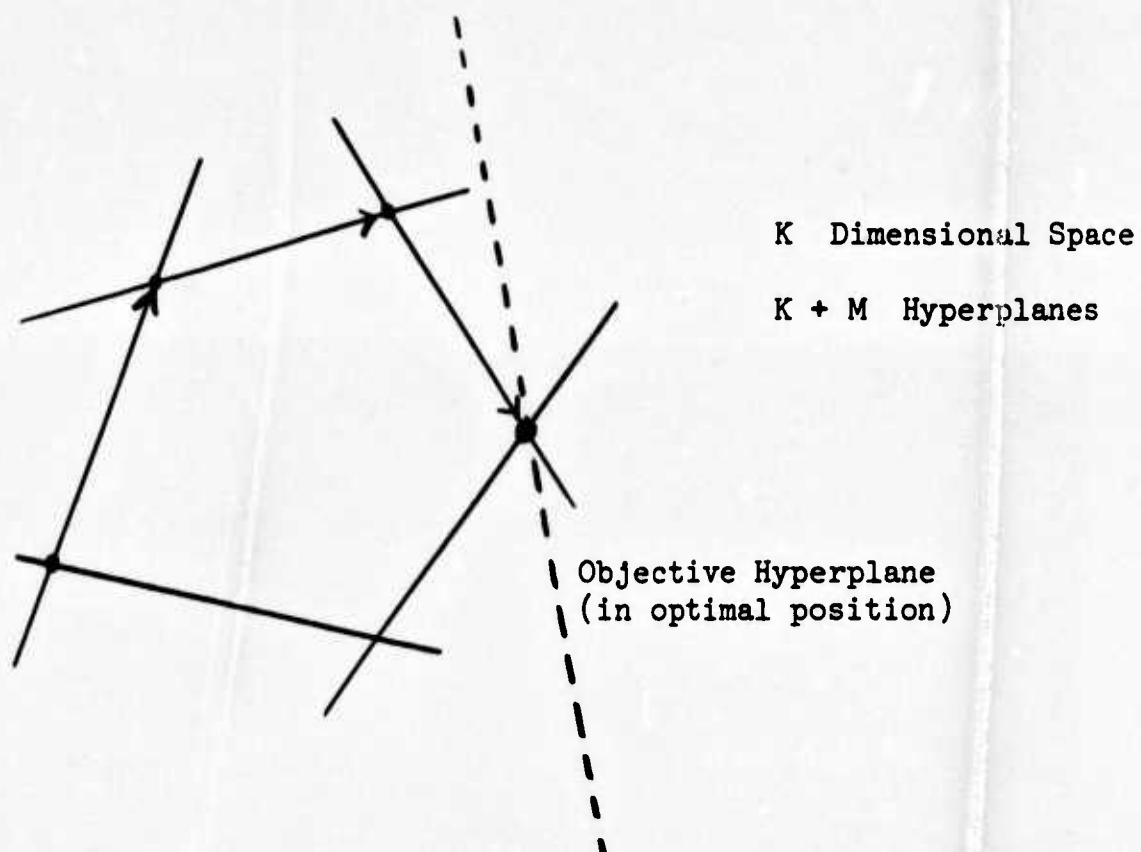


FIGURE 1

Problem 1: In the Simplex Method the number of vertex steps in practice is of the order of magnitude of  $M$  [the excess of the number of hyperplanes over the dimension]. Determine a sharp upper bound for the minimum number of steps.

solution. Now, it is a relatively easy matter to program a computer to pass from one extreme point to the next. One approach therefore, is to investigate all the extreme points and pick out the best. Unfortunately, some simple convex polytopes like an  $n$ -cube have a prohibitive number of vertices (the  $n$ -cube has  $2^n$  vertices) and this simple naive vertex-search approach is doomed to failure. Surprisingly enough, a minor variation of it has worked out quite well.

This method moves from one vertex to the next, moving along that edge which gives the greatest decrease in the objective function per unit change of the variable being introduced into the solution. This minor change, introducing a steepest descent, would not appear to be sufficient to prevent endless wandering over the edges of the polytope. Yet all over the world this procedure, called the Simplex Method, is a work-horse that daily solves thousands of problems. This brings us to the first of several famous problems of mathematical programming, which is stated below Figure 1. In practice, the number of steps is often of the magnitude of the excess  $m$  of the number of hyper-planes over the dimension of the space  $k$ . The open question is why?

The study of the properties of convex polytopes forms an important part of linear programming research. It is hoped that insight may be provided on why the simplex method is as efficient as it is and how it might be improved. A second unsolved problem, closely related to this one is the famous "m-step" or "Hirsch Conjecture" (see Figure 2).

Mathematics as we know it today is an eerie edifice of abstraction that has its roots in problems of the real world. But these problems, aside from puzzles and number lore, find their origins in the world of

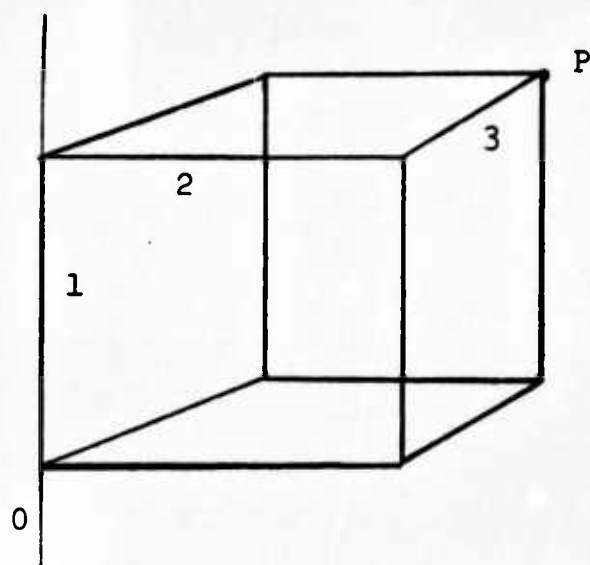


FIGURE 2

Problem 2: Given a convex polytope in  $M$  dimensional space with  $2M$  faces,  $M$  faces intersecting at  $0$ , the other  $M$  faces intersecting at  $P$  :

Prove (or Disprove): There exists a chain of  $M$  edges connecting  $0$   $T \subset P$  .

physics. The every day real world of planning and decisions attracted slight interest among mathematicians until about 1947. The reasons for this are obscure. In 1936 Motzkin, in his thesis, counted, after exhaustive search, some 30-odd papers on linear inequalities. [Compare this with the number that were published in a single month at that time on linear equations and the closely related matrix theory!] Those papers that were published were isolated one from the other. It is strange that there was so little interest, for it is easy to set up mathematical models of decision problems in the form of linear inequalities. Perhaps it might be argued that the reason such problems were never studied was the difficulty of solution without electronic computers. But why then were equation systems studied when even a 20x20 system was also out of the reach of available computers? I do not know the answer to the question as to why inequality theory was not investigated more intensively. Today we know that study of inequality constrained systems is an interesting and exciting field, full of challenges, like any other field of mathematics.

Part of the motivational drive comes from the desire by industry to solve bigger and bigger systems. In refinery operations, an initial simple blending model, will in time be expanded until it includes the distillation and reforming units, the selection of crude oils, oil field production, and the distribution and storage of the multi-product output. Industries which make numerous products in a number of alternative plants and ship to many destinations, give rise to very large scale mathematical models. One such system has over 30,000 equations and over  $10^6$  non-negative variables! In spite of their size, this and other such problems have been successfully solved. The technique used is called



the decomposition principle. It is a special application of the Lagrange Multiplier approach which permits the breaking up of the original problem into a sequence of smaller linear programs.

Lagrange, as you may recall, made many contributions to mathematics from number theory to analysis. It is said that when the post held by Euler was vacated, Frederick the Great invited Lagrange with the words: "from the greatest king in Europe to the greatest mathematician in Europe." Lagrange, by contrast, was a modest man.

The general approach is best illustrated by considering a very broad class of non-linear inequality constrained extreme problems. We begin by first considering classic equality constrained problem.

Find  $x$  such that

$$\begin{aligned} f_0(x) &= \min & x &= (x_1, \dots, x_n) \\ f_1(x) &= 0 \\ &\vdots \\ f_m(x) &= 0 \end{aligned}$$

We form the Lagrangian:

$$\phi(x) = f_0(x) + \lambda_1 f_1(x) + \dots + \lambda_m f_m(x)$$

Theorem: If  $x = \hat{x}$  yields  $\phi(\hat{x}) = \min \phi(x)$  and if (by good luck)  $f_i(\hat{x}) = 0$  then  $x$  solves the original problem.

For inequality constrained problems we have an analogous statement. We make the obvious change and, in addition require

$$\lambda_i f_i(\hat{x}) = 0, \quad \lambda_i \geq 0.$$

Lagrange's great contribution was to make "the good luck" happen by treating the constants as unknowns and choosing their values by back substitution so as to satisfy the restrictions. However, in certain applica-



tions, like linear programming, the Lagrange approach degenerates in an embarrassing manner that does not make it a practical approach. The basic theorem for inequalities constraints is due to Fritz John. Assuming  $f_i(x)$  are continuous and differentiable in the neighborhood of an extremal solution, he showed that the required multipliers exist if a multiplier  $\lambda_0$  for the  $f_0(x)$  function is included also. Kuhn-Tucker, assuming conditions that rule out such things as cusps, proved that a set of multipliers exist with  $\lambda_0 = 1$ .

We shall make a trivial extension of the Lagrangian approach that will have important implications for applications.

First Idea: Use a Partial Set of Lagrange Multipliers

$$\text{Min } f_0(x)$$

$$f_1(x) \leq 0 \quad : \quad \lambda_1$$

$$f_2(x) \leq 0 \quad : \quad \lambda_2$$

----- ←partition

$$f_3(x) \leq 0$$

$$f_4(x) \leq 0$$

sub-problem

$$\phi(x) = f_0(x) + \lambda_1 f_1(x) + \lambda_2 f_2(x)$$

Theorem: If  $x = \hat{x}$  yields  $\text{Min } \phi(x)$  subject to  $f_3(x) \leq 0$ ,  $f_4(x) \leq 0$  and if (by good luck)  $f_1(\hat{x}) \leq 0$ ,  $f_2(\hat{x}) \leq 0$ , then  $\hat{x}$  solves the original problem.

Lagrange's idea was to make the "good luck" happen

treat  $\lambda_i$  unknown constants

find  $\hat{x} = X(\lambda_1, \lambda_2)$

choose  $\lambda_1, \lambda_2$  so that  $f_1(\hat{x}) \leq 0$ ,  $f_2(\hat{x}) \leq 0$

For inequality systems, it is usually difficult to treat  $\hat{x}$  analytically as

a function of  $\lambda$ . To get around this difficulty, we try an alternate approach.

Second Idea: Start with a guess. For  $\lambda$ , say  $\lambda = \lambda^0$ , examine  $x = x_0$ , the solution to the sub-problem. Use it to obtain a better guess for  $\lambda = \lambda^0$ ; iterate.

Example: Find  $x \geq 0$ , Min  $f_0(x)$  :

$$c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5 = f_0(x)$$

Guess  $\lambda_1$ :  $a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 + a_{15}x_5 \leq b_2$

----- + partition

$$a_{21}x_1 + a_{22}x_2 \leq b_2$$

$$a_{31}x_1 + a_{32}x_2 \leq b_3$$

$$a_{43}x_3 + a_{44}x_4 \leq b_4$$

$$a_{53}x_3 + a_{54}x_4 \leq b_5$$

$$+ a_{65}x_5 \leq b_6$$

Sub-  
Problem

$$\phi_1(x_1, x_2) + \phi_2(x_3, x_4) + \phi_3(x_5) = \phi(x)$$

Subproblem: Decomposes into 3 separate minimization problems. Hence the term: "Decomposition Principle".

The decomposition algorithm may be described by a chicken-and-egg process. The "egg" part begins with an estimated value for  $\lambda_1 = \lambda_1^k$  and  $\lambda_2 = \lambda_2^k$ . This defines a Lagrangian problem with partial constraints. We assume we have some way to solve this and that  $x = x^k$ . Now  $x^k$  is used to initiate the "chicken" part by forming the vector  $[f_0(x^k), f_1(x^k), f_2(x^k), 1]$ . Earlier generations of chickens and eggs have already produced the other columns shown below in the "Master Linear Program". A convex combination of these columns is sought such

that their weighted average, component by component, is non-positive and the first component is minimal. This is a linear program. The Lagrange Multipliers associated with this program are called dual variables. Their optimal values, denoted by  $(1, \lambda_1, \lambda_2, \Delta) = (1, \lambda_1^{k+1}, \lambda_2^{k+1}, \Delta^{k+1})$ , is now used to initiate the next Lagrangian problem.

---

Master Linear Program: Find Min  $z$  and  $w_i \geq 0$  ; optimal dual variables  $(1, \lambda_1, \lambda_2, \Delta) = (1, \lambda_1^{k+1}, \lambda_2^{k+1}, \Delta^{k+1})$ .

$$\begin{array}{c|c} 1 & \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \\ \lambda_1 & \\ \lambda_2 & \\ \Delta & [0] \end{array} \quad z + \begin{array}{c} \begin{bmatrix} f_0(x^0) \\ f_1(x^0) \\ f_2(x^0) \end{bmatrix} \\ w_0 + \begin{bmatrix} f_0(x^1) \\ f_1(x^1) \\ f_2(x^2) \end{bmatrix} w_1 + \dots + \begin{bmatrix} f_0(x^k) \\ f_1(x^k) \\ f_2(x^k) \end{bmatrix} w_k \end{array} \leq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{array}{ccccccc} & & [1] & & [1] & & [1] \\ & & & & & & = [1] \end{array}$$


---

The proof that the iterative process converges assumes that the functions  $f_i(x)$  are convex. Applied to the earlier referenced large scale problem, the Master Program involved about 300 equations. The remainder, treated as one subproblem, had 30,000 mutually exclusive equations and therefore would break up into 30,000 independent problems. Applied to calculus of variations type problems, such as the optimal linear control processes considered by Pontryagin, Bellman, Zadeh and others, this approach can lead to a constructive solution technique.

Generally speaking, Lagrange Multipliers are treated in the literature as just a device. Courant describes it as an "elegant device" for getting around eliminating dependent variables and then solving a simple minimization problem without restrictions. Recent research, however, reveals that it has a deeper meaning.